

CRF Errors Corrected by the STIC Systems Branch

AU 1894
#20
10/28/93

Serial Number: 923,692C

CRF Processing Date: 10/28/93
 Edited by: mf
 Verified by: mf (STIC staff)

ENTERED

- ☐ Changed a file from non-ASCII to ASCII
- ☐ Changed the margins in cases where the sequence text was "wrapped" down to the next line
- ☐ Edited a format error in the Current Application Data section, specifically: _____
- ☐ Edited the Current Application Data section with the actual current number. The number inputted by the applicant was ☐ the prior application data; or ☐ other _____.
- ☐ Added the mandatory heading and subheadings for "Current Application Data".
- ☐ Edited the "Number of Sequences" field. The applicant spelled out a number instead of using an integer.
- ☐ Changed the spelling of a mandatory field (the headings or subheadings), specifically: _____
- ☐ Inserted a space between the last nucleic designator and the nucleic number for sequences: _____
- ☐ Deleted page numbers in the text of the sequence listing, which is considered invalid text.
- ☐ Corrected the SEQ ID NO when obviously incorrect. The sequence numbers that were edited were: _____
- ☐ Inserted a nucleic number at the end of a nucleic line. SEQ ID NO's edited: _____
- ☐ Corrected subheading placement. All responses must be on the same line as each subheading. If the applicant placed a response below the subheading, this was moved to its appropriate place.
- ☐ Inserted colons after headings/subheadings. Headings edited included: _____
- ☐ Deleted extra, invalid, headings used by an applicant, specifically: _____
- ☐ Deleted non-ASCII "garbage" at the end of files, and other invalid text, such as a secretary's initials.
- ☐ Inserted mandatory headings, specifically: _____
- ☒ Corrected an obvious error in the response, specifically:
changed a letter "o" to a number zero (p. 12)
- ☐ Edited identifiers where upper case is used but lower case is required, or vice versa.
- ☐ Corrected an error in the Number of Sequences field, specifically: _____
- ☐ A "Hard Page Break" code was inserted by the applicant. All occurrences had to be deleted.
- ☐ Other: _____

***Examiner: The above corrections must be communicated to the applicant in the first Office Action. DO NOT send a copy of this form.**

8/01/93

RAW SEQUENCE LISTING
PATENT APPLICATION US/07/923,692CDATE: 10/28/93
TIME: 16:28:02

INPUT SET: S818.raw

SEQUENCE LISTING

ENTERED

1
2
3 (1) General Information:
4
5 (i) APPLICANT: Donson, Jon
6 Dawson, William O.
7 Grantham, George L.
8 Turpen, Thomas H.
9 Turpen, Ann Myers
10 Garger, Stephen J.
11 Grill, Laurence K.
12
13 (ii) TITLE OF INVENTION: RECOMBINANT PLANT VIRAL NUCLEIC ACIDS
14
15 (iii) NUMBER OF SEQUENCES: 11
16
17 (iv) CORRESPONDENCE ADDRESS:
18 (A) ADDRESSEE: Limbach & Limbach
19 (B) STREET: 2001 Ferry Building
20 (C) CITY: San Francisco
21 (D) STATE: CAL
22 (F) ZIP: 94111
23
24 (v) COMPUTER READABLE FORM:
25 (A) MEDIUM TYPE: Floppy disk
26 (B) COMPUTER: IBM PC compatible
27 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
28 (D) SOFTWARE: Patent in Release #1.0, Version #1.25
29
30 (vi) CURRENT APPLICATION DATA:
31 (A) APPLICATION NUMBER: US 923,692
32 (B) FILING DATE: 31-JUL-1992
33 (C) CLASSIFICATION:
34 (vii) PRIOR APPLICATION DATA:
35 (A) APPLICATION NUMBER: US 600,244
36 (B) FILING DATE: 22-OCT-1990
37
38 (vii) PRIOR APPLICATION DATA:
39 (A) APPLICATION NUMBER: US 641,617
40 (B) FILING DATE: 16-JAN-1991
41
42 (vii) PRIOR APPLICATION DATA:
43 (A) APPLICATION NUMBER: US 310,881
44 (B) FILING DATE: 17-FEB-1989
45
46 (vii) PRIOR APPLICATION DATA:
47 (A) APPLICATION NUMBER: US 160,766
48 (B) FILING DATE: 26-FEB-1988
49
50 (vii) PRIOR APPLICATION DATA:
51 (A) APPLICATION NUMBER: US 160,771

RAW SEQUENCE LISTING
PATENT APPLICATION US/07/923,692CDATE: 10/28/93
TIME: 16:28:05

INPUT SET: S818.raw

52 (B) FILING DATE: 26-FEB-1988
53
54 (vii) PRIOR APPLICATION DATA:
55 (A) APPLICATION NUMBER: US 347,637
56 (B) FILING DATE: 05-MAY-1989
57
58 (vii) PRIOR APPLICATION DATA:
59 (A) APPLICATION NUMBER: US 363,138
60 (B) FILING DATE: 08-JUN-1989
61
62 (vii) PRIOR APPLICATION DATA:
63 (A) APPLICATION NUMBER: US 219,279
64 (B) FILING DATE: 15-JUL-1988
65
66 (viii) ATTORNEY/AGENT INFORMATION:
67 (A) NAME: Halluin, Albert P.
68 (B) REGISTRATION NUMBER: 28,957
69 (C) REFERENCE/DOCKET NUMBER: BIOG-20121
70 USA
71
72 (ix) TELECOMMUNICATION INFORMATION:
73 (A) TELEPHONE: 415-433-4150
74 (B) TELEFAX: 415-433-8716
75
76
77 (2) INFORMATION FOR SEQ ID NO: 1:
78
79 (i) SEQUENCE CHARACTERISTICS:
80 (A) LENGTH: 4 amino acids
81 (B) TYPE: amino acid
82 (D) TOPOLOGY: linear
83
84 (ii) MOLECULE TYPE: peptide
85
86 (iii) HYPOTHETICAL: NO
87
88 (iv) ANTI-SENSE: NO
89
90 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
91
92 Pro Xaa Gly Pro
93 1
94
95 (2) INFORMATION FOR SEQ ID NO: 2:
96
97 (i) SEQUENCE CHARACTERISTICS:
98 (A) LENGTH: 13 base pairs
99 (B) TYPE: nucleic acid
100 (C) STRANDEDNESS: single
101 (D) TOPOLOGY: linear
102

RAW SEQUENCE LISTING
PATENT APPLICATION US/07/923,692CDATE: 10/28/93
TIME: 16:28:07

INPUT SET: S818.raw

103 (ii) MOLECULE TYPE: DNA (genomic)
104
105 (iii) HYPOTHETICAL: NO
106
107 (iv) ANTI-SENSE: NO
108
109 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
110
111 GGGTACCTGG GCC 13
112
113
114
115 (2) INFORMATION FOR SEQ ID NO: 3:
116
117 (i) SEQUENCE CHARACTERISTICS:
118 (A) LENGTH: 886 base pairs
119 (B) TYPE: nucleic acid
120 (C) STRANDEDNESS: single
121 (D) TOPOLOGY: linear
122
123 (ii) MOLECULE TYPE: DNA (genomic)
124
125 (iii) HYPOTHETICAL: NO
126
127 (iv) ANTI-SENSE: NO
128
129 (vi) ORIGINAL SOURCE:
130 (A) ORGANISM: Chinese cucumber
131
132 (vii) IMMEDIATE SOURCE:
133 (B) CLONE: alpha-trichosanthin
134
135 (ix) FEATURE:
136 (A) NAME/KEY: CDS (B) LOCATION: 8. .877
137 (B) LOCATION: 8. .877
138
139 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
140
141 CTCGAGG ATG ATC AGA TTC TTA GTC CTC TCT TTG CTA ATT CTC ACC CTC 49
142
143 Met Ile Arg Phe Leu Val Leu Ser Leu Leu Ile Leu Thr Leu
144 1 5 10
145
146 TTC CTA ACA ACT CCT GCT GTG GAG GGC GAT GTT AGC TTC CGT TTA TCA 97
147
148 Phe Leu Thr Thr Pro Ala Val Glu Gly Asp Val Ser Phe Arg Leu Ser
149 15 20 25 30
150
151 GGT GCA ACA AGC AGT TCC TAT GGA GTT TTC ATT TCA AAT CTG AGA AAA 145
152
153 Gly Ala Thr Ser Ser Ser Tyr Gly Val Phe Ile Ser Asn Leu Arg Lys

RAW SEQUENCE LISTING PATENT APPLICATION US/07/923,692C

DATE: 10/28/93
TIME: 16:28:10

INPUT SET: S818.raw

154		35		40		45		
155								
156	GCT CTT CCA AAT GAA AGG AAA CTG TAC GAT ATC CCT CTG TTA CGT TCC							193
157								
158	Ala Leu Pro Asn Glu Arg Lys Leu Tyr Asp Ile Pro Leu Leu Arg Ser							
159		50		55		60		
160								
161	TCT CTT CCA GGT TCT CAA CGC TAC GCA TTG ATC CAT CTC ACA AAT TAC							241
162								
163	Ser Leu Pro Gly Ser Gln Arg Tyr Ala Leu Ile His Leu Thr Asn Tyr							
164		65		70		75		
165								
166	GCC GAT GAA ACC ATT TCA GTG GCC ATA GAC GTA ACG AAC GTC TAT ATT							289
167								
168	Ala Asp Glu Thr Ile Ser Val Ala Ile Asp Val Thr Asn Val Tyr Ile							
169		80		85		90		
170								
171	ATG GGA TAT CGC GCT GGC GAT ACA TCC TAT TTT TTC AAC GAG GCT TCT							337
172								
173	Met Gly Tyr Arg Ala Gly Asp Thr Ser Tyr Phe Phe Asn Glu Ala Ser							
174		95		100		105		110
175								
176	GCA ACA GAA GCT GCA AAA TAT GTA TTC AAA GAC GCT ATG CGA AAA GTT							385
177								
178	Ala Thr Glu Ala Ala Lys Tyr Val Phe Lys Asp Ala Met Arg Lys Val							
179		115		120		125		
180								
181	ACG CTT CCA TAT TCT GGC AAT TAC GAA AGG CTT CAA ACT GCT GCG GGC							433
182								
183	Thr Leu Pro Tyr Ser Gly Asn Tyr Glu Arg Leu Gln Thr Ala Ala Gly							
184		130		135		140		
185								
186	AAA ATA AGG GAA AAT ATT CCG CTT GGA CTC CCA GCT TTG GAC AGT GCC							481
187								
188	Lys Ile Arg Glu Asn Ile Pro Leu Gly Leu Pro Ala Leu Asp Ser Ala							
189		145		150		155		
190								
191	ATT ACC ACT TTG TTT TAC TAC AAC GCC AAT TCT GCT GCG TCG GCA CTT							529
192								
193	Ile Thr Thr Leu Phe Tyr Tyr Asn Ala Asn Ser Ala Ala Ser Ala Leu							
194		160		165		170		
195								
196	ATG GTA CTC ATT CAG TCG ACG TCT GAG GCT GCG AGG TAT AAA TTT ATT							577
197								
198	Met Val Leu Ile Gln Ser Thr Ser Glu Ala Ala Arg Tyr Lys Phe Ile							
199		175		180		185		190
200								
201	GAG CAA CAA ATT GGG AAG CGC GTT GAC AAA ACC TTC CTA CCA AGT TTA							625
202								
203	Glu Gln Gln Ile Gly Lys Arg Val Asp Lys Thr Phe Leu Pro Ser Leu							
204		195		200		205		

RAW SEQUENCE LISTING
PATENT APPLICATION US/07/923,692CDATE: 10/28/93
TIME: 16:28:12

INPUT SET: S818.raw

205
206 GCA ATT ATA AGT TTG GAA AAT AGT TGG TCT GCT CTC TCC AAG CAA ATT 673
207
208 Ala Ile Ile Ser Leu Glu Asn Ser Trp Ser Ala Leu Ser Lys Gln Ile
209 210 215 220
210
211 CAG ATA GCG AGT ACT AAT AAT GGA CAG TTT GAA ACT CCT GTT GTG CTT 721
212
213 Gln Ile Ala Ser Thr Asn Asn Gly Gln Phe Glu Thr Pro Val Val Leu
214 225 230 235
215
216 ATA AAT GCT CAA AAC CAA CGA GTC ATG ATA ACC AAT GTT GAT GCT GGA 769
217
218 Ile Asn Ala Gln Asn Gln Arg Val Met Ile Thr Asn Val Asp Ala Gly
219 240 245 250
220
221 GTT GTA ACC TCC AAC ATC GCG TTG CTG CTG AAT CGA AAC AAT ATG GCA 817
222
223 Val Val Thr Ser Asn Ile Ala Leu Leu Leu Asn Arg Asn Asn Met Ala
224 255 260 265 270
225
226 GCC ATG GAT GAC GAT GTT CCT ATG ACA CAG AGC TTT GGA TGT GGA AGT 865
227
228 Ala Met Asp Asp Asp Val Pro Met Thr Gln Ser Phe Gly Cys Gly Ser
229 275 280 285
230
231 TAT GCT ATT TAGTAACTCG AG 886
232
233 Tyr Ala Ile
234 290
235
236
237 (2) INFORMATION FOR SEQ ID NO:4:
238
239 (i) SEQUENCE CHARACTERISTICS:
240 (A) LENGTH: 289 amino acids
241 (B) TYPE: amino acid
242 (D) TOPOLOGY: linear
243
244 (ii) MOLECULE TYPE: protein
245
246 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:
247
248
249 Met Ile Arg Phe Leu Val Leu Ser Leu Leu Ile Leu Thr Leu Phe Leu
250 1 5 10 15
251
252 Thr Thr Pro Ala Val Glu Gly Asp Val Ser Phe Arg Leu Ser Gly Ala
253 20 25 30
254
255 Thr Ser Ser Ser Tyr Gly Val Phe Ile Ser Asn Leu Arg Lys Ala Leu

RAW SEQUENCE LISTING PATENT APPLICATION US/07/923,692C

DATE: 10/28/93
TIME: 16:28:14

INPUT SET: S818.raw

```

256          35          40          45
257
258 Pro Asn Glu Arg Lys Leu Tyr Asp Ile Pro Leu Leu Arg Ser Ser Leu
259      50          55          60
260
261 Pro Gly Ser Gln Arg Tyr Ala Leu Ile His Leu Thr Asn Tyr Ala Asp
262      65          70          75          80
263
264 Glu Thr Ile Ser Val Ala Ile Asp Val Thr Asn Val Tyr Ile Met Gly
265          85          90          95
266
267 Tyr Arg Ala Gly Asp Thr Ser Tyr Phe Phe Asn Glu Ala Ser Ala Thr
268          100          105          110
269
270 Glu Ala Ala Lys Tyr Val Phe Lys Asp Ala Met Arg Lys Val Thr Leu
271          115          120          125
272
273 Pro Tyr Ser Gly Asn Tyr Glu Arg Leu Gln Thr Ala Ala Gly Lys Ile
274          130          135          140
275
276 Arg Glu Asn Ile Pro Leu Gly Leu Pro Ala Leu Asp Ser Ala Ile Thr
277 145          150          155          160
278
279 Thr Leu Phe Tyr Tyr Asn Ala Asn Ser Ala Ala Ser Ala Leu Met Val
280          165          170          175
281
282 Leu Ile Gln Ser Thr Ser Glu Ala Ala Arg Tyr Lys Phe Ile Glu Gln
283          180          185          190
284
285 Gln Ile Gly Lys Arg Val Asp Lys Thr Phe Leu Pro Ser Leu Ala Ile
286          195          200          205
287
288 Ile Ser Leu Glu Asn Ser Trp Ser Ala Leu Ser Lys Gln Ile Gln Ile
289          210          215          220
290
291 Ala Ser Thr Asn Asn Gly Gln Phe Glu Thr Pro Val Val Leu Ile Asn
292 225          230          235          240
293
294 Ala Gln Asn Gln Arg Val Met Ile Thr Asn Val Asp Ala Gly Val Val
295          245          250          255
296
297 Thr Ser Asn Ile Ala Leu Leu Leu Asn Arg Asn Asn Met Ala Ala Met
298          260          265          270
299
300 Asp Asp Asp Val Pro Met Thr Gln Ser Phe Gly Cys Gly Ser Tyr Ala
301          275          280          285
302 Ile
303
304
305 (2) INFORMATION FOR SEQ ID NO: 5:
306

```

RAW SEQUENCE LISTING PATENT APPLICATION US/07/923,692C

DATE: 10/28/93
TIME: 16:28:16

INPUT SET: S818.raw

```

307      (i)      SEQUENCE CHARACTERISTICS:
308                (A)  LENGTH: 1450 base pairs
309                (B)  TYPE: nucleic acid
310                (C)  STRANDEDNESS: single
311                (D)  TOPOLOGY: linear
312
313      (ii)      MOLECULE TYPE: DNA (genomic)
314
315      (iii)      HYPOTHETICAL: NO
316
317      (iv)      ANTI-SENSE: NO
318
319      (vi)      ORIGINAL SOURCE:
320                (A)  ORGANISM: Oryza sativa
321
322      (vii)      IMMEDIATE SOURCE:
323                (B)  CLONE: alpha-amylase
324
325      (ix)      FEATURE:
326                (A)  NAME/KEY: CDS (B) LOCATION: 12. .1316
327                (B)  LOCATION: 12. .1316
328
329      (xi)      SEQUENCE DESCRIPTION: SEQ ID NO: 5:
330
331      CCTCGAGGTG C ATG CAG GTG CTG AAC ACC ATG GTG AAC A CAC TTC TTG      48
332
333                Met Gln Val Leu Asn Thr Met Val Asn Lys His Phe Leu
334                1                5                10
335
336      TCC CTT TCG GTC CTC ATC GTC CTC CTT GGC CTC TCC TCC AAC TTG ACA      96
337
338      Ser Leu Ser Val Leu Ile Val Leu Leu Gly Leu Ser Ser Asn Leu Thr
339                15                20                25
340
341      GCC GGG CAA GTC CTG TTT CAG GGA TTC AAC TGG GAG TCG TGG AAG GAG      144
342
343      Ala Gly Gln Val Leu Phe Gln Gly Phe Asn Trp Glu Ser Trp Lys Glu
344                30                35                40                45
345
346      AAT GGC GGG TGG TAC AAC TTC CTG ATG GGC AAG GTG GAC GAC ATC GCC      192
347
348      Asn Gly Gly Trp Tyr Asn Phe Leu Met Gly Lys Val Asp Asp Ile Ala
349                50                55                60
350
351      GCA GCC GGC ATC ACC CAC GTC TGG CTC CCT CCG CCG TCT CAC TCT GTC      240
352
353      Ala Ala Gly Ile Thr His Val Trp Leu Pro Pro Pro Ser His Ser Val
354                65                70                75
355
356      GGC GAG CAA GGC TAC ATG CCT GGG CGG CTG TAC GAT CTG GAC GCG TCT      288
357

```


RAW SEQUENCE LISTING PATENT APPLICATION US/07/923,692C

DATE: 10/28/93
TIME: 16:28:29

INPUT SET: S818.raw

358	Gly	Glu	Gln	Gly	Tyr	Met	Pro	Gly	Arg	Leu	Tyr	Asp	Leu	Asp	Ala	Ser	
359			80					85					90				
360																	
361	AAG	TAC	GGC	AAC	GAG	GCG	CAG	CTC	AAG	TCG	CTG	ATC	GAG	GCG	TTC	CAT	336
362																	
363	Lys	Tyr	Gly	Asn	Glu	Ala	Gln	Leu	Lys	Ser	Leu	Ile	Glu	Ala	Phe	His	
364		95					100					105					
365																	
366	GGC	AAG	GGC	GTC	CAG	GTG	ATC	GCC	GAC	ATC	GTC	ATC	AAC	CAC	CGC	ACG	384
367																	
368	Gly	Lys	Gly	Val	Gln	Val	Ile	Ala	Asp	Ile	Val	Ile	Asn	His	Arg	Thr	
369	110				115						120					125	
370																	
371	GCG	GAG	CAC	AAG	GAC	GGC	CGC	GGC	ATC	TAC	TGC	CTC	TTC	GAG	GGC	GGG	432
372																	
373	Ala	Glu	His	Lys	Asp	Gly	Arg	Gly	Ile	Tyr	Cys	Leu	Phe	Glu	Gly	Gly	
374					130					135					140		
375																	
376	ACG	CCC	GAC	TCC	CGC	CTC	GAC	TGG	GGC	CCG	CAC	ATG	ATC	TGC	CGC	GAC	480
377																	
378	Thr	Pro	Asp	Ser	Arg	Leu	Asp	Trp	Gly	Pro	His	Met	Ile	Cys	Arg	Asp	
379				145					150					155			
380																	
381	GAC	CCC	TAC	GGC	CAT	GGC	ACC	GGC	AAC	CCG	GAC	ACC	GGC	GCC	GAC	TTC	528
382																	
383	Asp	Pro	Tyr	Gly	Asp	Gly	Thr	Gly	Asn	Pro	Asp	Thr	Gly	Ala	Asp	Phe	
384			160					165					170				
385																	
386	GCC	GCC	GCG	CCG	GAC	ATC	GAC	CAC	CTC	AAC	AAG	CGC	GTC	CAG	CGG	GAG	576
387																	
388	Ala	Ala	Ala	Pro	Asp	Ile	Asp	His	Leu	Asn	Lys	Arg	Val	Gln	Arg	Glu	
389		175					180					185					
390																	
391	CTC	ATT	GGC	TGG	CTC	GAC	TGG	CTC	AAG	ATG	GAC	ATC	GGC	TTC	GAC	GCG	624
392																	
393	Leu	Ile	Gly	Trp	Leu	Asp	Trp	Leu	Lys	Met	Asp	Ile	Gly	Phe	Asp	Ala	
394	190					195					200					205	
395																	
396	TGG	CGC	CTC	GAC	TTC	GCC	AAG	GGC	TAC	TCC	GCC	GAC	ATG	GCA	AAC	ATC	672
397																	
398	Trp	Arg	Leu	Asp	Phe	Ala	Lys	Gly	Tyr	Ser	Ala	Asp	Met	Ala	Lys	Ile	
399				210						215					220		
400																	
401	TAC	ATC	GAC	GCC	ACC	GAG	CCG	AGC	TTC	GCC	GTG	CCC	GAG	ATA	TCG	ACG	720
402																	
403	Tyr	Ile	Asp	Ala	Thr	Glu	Pro	Ser	Phe	Ala	Val	Ala	Glu	Ile	Trp	Thr	
404				225					230					235			
405																	
406	TCC	ATG	GCG	AAC	GGC	GGG	GAC	GGC	AAG	CCG	AAC	TAC	GAC	CAG	AAC	GCG	768
407																	
408	Ser	Met	Ala	Asn	Gly	Gly	Asp	Gly	Lys	Pro	Asn	Tyr	Asp	Gln	Asn	Ala	

RAW SEQUENCE LISTING PATENT APPLICATION US/07/923,692C

DATE: 10/28/93
TIME: 16:28:36

INPUT SET: S818.raw

409		240		245		250												
410																		
411	CAC	CGG	CAG	GAG	CTG	GTC	AAC	TGG	GTC	GAT	CGT	GTC	GGC	GGC	GCC	AAC		816
412																		
413	His	Arg	Gln	Glu	Leu	Val	Asn	Trp	Val	Asp	Arg	Val	Gly	Gly	Ala	Asn		
414		255					260					265						
415																		
416	ACC	AAC	GGC	ACG	GCG	TTC	GAC	TTC	ACC	ACC	AAG	GGC	ATC	CTC	AAC	GTC		864
417																		
418	Ser	Asn	Gly	Thr	Ala	Phe	Asp	Phe	Thr	Thr	Lys	Gly	Ile	Leu	Asn	Val		
419		270				275					280				285			
420																		
421	GCC	GTG	GAG	GGC	GAG	CTG	TGG	CGC	CTC	CGC	GGC	GAG	GAC	GGC	AAG	GCG		912
422																		
423	Ala	Val	Glu	Gly	Glu	Leu	Trp	Arg	Leu	Arg	Gly	Glu	Asp	Gly	Lys	Ala		
424					290					295					300			
425																		
426	CCC	GGC	ATG	ATC	GGG	TGC	TGG	CCG	GCC	AAG	GCG	ACG	ACC	TTC	GTC	GAC		960
427																		
428	Pro	Gly	Met	Ile	Gly	Trp	Trp	Pro	Ala	Lys	Ala	Thr	Thr	Phe	Val	Asp		
429				305					310					315				
430																		
431	AAC	CAC	GAC	ACC	GGC	TCG	ACG	CAG	CAC	CTG	TGG	CCG	TTC	CCC	TCC	GAC		1008
432																		
433	Asn	His	Asp	Thr	Gly	Ser	Thr	Gln	His	Leu	Trp	Pro	Phe	Pro	Ser	Asp		
434			320					325					330					
435																		
436	AAG	GTC	ATG	CAG	GGC	TAC	GCA	TAC	ATC	CTC	ACC	CAC	CCC	GGC	AAC	CCA		1056
437																		
438	Lys	Val	Met	Gln	Gly	Tyr	Ala	Tyr	Ile	Leu	Thr	His	Pro	Gly	Asn	Pro		
439		335					340					345						
440																		
441	TGC	ATC	TTG	TAC	GAC	CAT	TTC	TTC	GAT	TGG	GGT	CTC	AAG	GAG	GAG	ATC		1104
442																		
443	Cys	Ile	Phe	Tyr	Asp	His	Phe	Phe	Asp	Trp	Gly	Leu	Lys	Glu	Glu	Ile		
444		350				355					360				365			
445																		
446	GAG	CGC	CTG	GTG	TCA	ATC	AGA	AAC	CGG	CAG	GGG	ATC	CAC	CCG	GCG	AGC		1152
447																		
448	Glu	Arg	Leu	Val	Ser	Ile	Arg	Asn	Arg	Gln	Gly	Ile	His	Pro	Ala	Ser		
449					370					375					380			
450																		
451	GAG	CTG	CGC	ATC	ATG	GAA	GCT	GAC	AGC	GAT	CTC	TAC	CTC	GCG	GAG	ATC		1200
452																		
453	Glu	Leu	Arg	Ile	Met	Glu	Ala	Asp	Ser	Asp	Leu	Tyr	Leu	Ala	Glu	Ile		
454				385					390					395				
455																		
456	GAT	GGC	AAG	GTG	ATC	ACA	AAG	ATT	GGA	CCA	AGA	TAC	GAC	GTC	GAA	CAC		1248
457																		
458	Asp	Gly	Lys	Val	Ile	Thr	Lys	Ile	Gly	Pro	Arg	Tyr	Asp	Val	Glu	His		
459			400					405					410					

RAW SEQUENCE LISTING PATENT APPLICATION US/07/923,692C

DATE: 10/28/93
TIME: 16:28:43

INPUT SET: S818.raw

```

460
461 CTC ATC CCC GAA GGC TTC CAG GTC GTC GCG CAC GGT GAT GGC TAC GCA      1296
462
463 Leu Ile Pro Glu Gly Phe Gln Val Val Ala His Gly Asp Gly Tyr Ala
464     415                      420                      425
465
466 ATC TGG GAG AAA ATC TGAGCGCACG ATGACGAGAC TCTCAGTTTA GCAGATTTAA      1351
467
468 Ile Trp Glu Lys LIe
469     430                      435
470
471 CCTGCGATTT TTACCCTGAC CGGTATACGT ATATACGTGC CGGCAACGAG CTGTATCCGA      1411
472
473
474 TCCGAATTAC GGATGCAATT GTCCACGAAG TCCTCGAGG      1450
475
476
477
478 (2) INFORMATION FOR SEQ ID NO: 6:
479
480 (i) SEQUENCE CHARACTERISTICS:
481 (A) LENGTH: 434 amino acids
482 (B) TYPE: amino acid
483 (D) Topology: linear
484
485 (ii) MOLECULE TYPE: protein
486
487 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
488
489 Met Gln Val Leu Asn Thr Met Val Asn Lys His Phe Leu Ser Leu Ser
490     1              5              10              15
491
492 Val Leu Ile Val Leu Leu Gly Leu Ser Ser Asn Leu Thr Ala Gly Gln
493     20              25              30
494
495 Val Leu Phe Gln Gly Phe Asn Trp Glu Ser Trp Lys Glu Asn Gly Gly
496     35              40              45
497
498 Trp Tyr Asn Phe Leu Met Gly Lys Val Asp Asp Ile Ala Ala Ala Gly
499     50              55              60
500
501 Ile Thr His Val Trp Leu Pro Pro Pro Ser His Ser Val Gly Glu Gln
502     65              70              75              80
503
504 Gly Tyr Met Pro Gly Arg Leu Tyr Asp Leu Asp Ala Ser Lys Tyr Gly
505     85              90              95
506
507 Asn Glu Ala Gln Leu Lys Ser Leu Ile Glu Ala Phe His Gly Lys Gly
508     100             105             110
509
510 Val Gln Val Ile Ala Asp Ile Val Ile Asn His Arg Thr Ala Glu His

```

RAW SEQUENCE LISTING PATENT APPLICATION US/07/923,692C

DATE: 10/28/93
TIME: 16:28:49

INPUT SET: S818.raw

```

511          115          120          125
512
513 Lys Asp Gly Arg Gly Ile Tyr Cys Leu Phe Glu Gly Gly Thr Pro Asp
514      130          135          140
515
516 Ser Arg Leu Asp Trp Gly Pro His Met Ile Cys Arg Asp Asp Pro Tyr
517 145          150          155          160
518
519 Gly Asp Gly Thr Gly Asn Pro Asp Thr Gly Ala Asp Phe Ala Ala Ala
520      165          170          175
521
522 Pro Asp Ile Asp His Leu Asn Lys Arg Val Gln Arg Glu Leu Ile Gly
523      180          185          190
524
525 Trp Leu Asp Trp Leu Lys Met Asp Ile Gly Phe Asp Ala Trp Arg Leu
526      195          200          205
527
528 Asp Phe Ala Lys Gly Tyr Ser Ala Asp Met Ala Lys Ile Tyr Ile Asp
529      210          215          220
530
531 Ala Thr Glu Pro Ser Phe Ala Val Ala Glu Ile Trp Thr Ser Met Ala
532 225          230          235          240
533
534 Asn Gly Gly Asp Gly Lys Pro Asn Tyr Asp Gln Asn Ala His Arg Gln
535      245          250          255
536
537 Glu Leu Val Asn Trp Val Asp Arg Val Gly Gly Ala Asn Ser Asn Gly
538      260          265          270
539
540 Thr Ala Phe Asp Phe Thr Thr Lys Gly Ile Leu Asn Val Ala Val Glu
541      275          280          285
542
543 Gly Glu Leu Trp Arg Leu Arg Gly Glu Asp Gly Lys Ala Pro Gly Met
544 290          295          300
545
546 Ile Gly Trp Trp Pro Ala Lys Ala Thr Thr Phe Val Asp Asn His Asp
547 305          310          315          320
548
549 Thr Gly Ser Thr Gln His Leu Trp Pro Phe Pro Ser Asp Lys Val Met
550      325          330          335
551
552 Gln Gly Tyr Ala Tyr Ile Leu Thr His Pro Gly Asn Pro Cys Ile Phe
553      340          345          350
554
555 Tyr Asp His Phe Phe Asp Trp Gly Leu Lys Glu Glu Ile Glu Arg Leu
556      355          360          365
557
558 Val Ser Ile Arg Asn Arg Gln Gly Ile His Pro Ala Ser Glu Leu Arg
559      370          375          380
560
561 Ile Met Glu Ala Asp Ser Asp Leu Tyr Leu Ala Glu Ile Asp Gly Lys

```

RAW SEQUENCE LISTING
PATENT APPLICATION US/07/923,692CDATE: 10/28/93
TIME: 16:28:56

INPUT SET: S818.raw

562 385 390 395 400
563
564 Val Ile Thr Lys Ile Gly Pro Arg Tyr Asp Val Glu His Leu Ile Pro
565 405 410 415
566
567 Glu Gly Phe Gln Val Val Ala His Gly Asp Gly Tyr Ala Ile Trp Glu
568 420 425 430
569
570 Lys Ile
571
572

573 (2) INFORMATION FOR SEQ ID NO:7:

574 (i) SEQUENCE CHARACTERISTICS:

575 (A) LENGTH: 709 base pairs

576 (B) TYPE: nucleic acid

577 (G) STRANDEDNESS: single

578 (D) TOPOLOGY: linear

579

580 (ii) MOLECULE TYPE: cDNA to mRNA

581

582 (iii) HYPOTHETICAL: NO

583

584 (iv) ANTI-SENSE: NO

585

586 (vi) ORIGINAL SOURCE:

587 (A) ORGANISM: Homo sapiens

588

589 (vii) IMMEDIATE SOURCE:

590 (B) CLONE: alpha-hemoglobin

591

592

593 (ix) FEATURE:

594 (A) NAME/KEY: transit_peptide (B)

595 LOCATION: 26. .241

596 (B) LOCATION: 26. .241

597

598 (ix) FEATURE:

599 (A) NAME/KEY: CDS

600 (B) LOCATION: 245. .670

601

602 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

603

604 CTCGAGGGCA TCTGATCTTT CAAGAATGGC ACAAATTAAC AACATGGCAC AAGGGATACA 60

605

606 AACCCCTTAAT CCCAATTCCA ATTTCCATAA ACCCCAAGTT CCTAAATCTT CAAGTTTCTTCT 120

607

608 TGTTTTTGGG TGTAACAAAC TGAAAAATTC AGCAAATCTT ATGTTGGTTT TGAAAAAAGA 180

609

610 TTCAATTTTT ATGCAAAAGT TTTGTTCCCTT TAGGATTTCG GCAGGTGGTA GAGTTTCTTG 240

611

612 CATG GTG CTG TCT CCT GCC GAC AAG ACC AAC GTC AAG GCC GCC TGG GGC 289

RAW SEQUENCE LISTING PATENT APPLICATION US/07/923,692C

DATE: 10/28/93
TIME: 16:29:03

INPUT SET: S818.raw

```

613
614      Val Leu Ser Pro Ala Asp Lys Thr Asn Val Lys Ala Ala Trp Cly
615      1              5              10              15
616
617  AAG GTT GGC GCG CAC GCT GGC GAG TAT GGT GCG GAG GCC CTG GAG AGG      337
618
619  Lys Val Gly Ala His Ala Gly Glu Tyr Gly Ala Glu Ala Leu Glu Arg
620      20              25              30
621
622  ATG TTC CTG TCC TTC CCC ACC ACC AAG ACC TAC TTC CCG CAC TTC GAC      385
623
624  Met Phe Leu Ser Phe Pro Thr Thr Lys Thr Tyr Phe Pro His Phe Asp
625      35              40              45
626
627  CTG AGC CAC GGC TCT GCC CAG GTT AAG GGC CAC GGC AAG AAG GTG GCC      433
628
629  Leu Ser His Gly Ser Ala Gln Val Lys Gly His Gly Lys Lys Val Ala
630      50              55              60
631
632  GAC GCG CTG ACC AAC GCC GTG GCG CAC GTG GAC GAC ATG CCC AAC GCG      481
633
634  Asp Ala Leu Thr Asn Ala Val Ala His Val Asp Asp Met Pro Asn Ala
635      65              70              75
636
637  CTG TCC GCC CTG AGC GAC CTG CAC GCG CAC AAG CTT CGG GTG GAC CCG      529
638
639  Leu Ser Ala Leu Ser Asp Leu His Ala His Lys Leu Arg Val Asp Pro
640      80              85              90              95
641
642  GTC AAC TTC AAG CTC CTA AGC CAC TGC CTG CTG GTG ACC CTG GCC GCC      577
643
644  Val Asn Phe Lys Leu Leu Ser His Cys Leu Leu Val Thr Leu Ala Ala
645      100              105              110
646
647  CAC CTC CCC GCC GAG TTC ACC CCT GCG GTG CAC GCC TCC CTG GAC AAG      625
648
649  His Leu Pro Ala Glu Phe Thr Pro Ala Val His Ala Ser Leu Asp Lys
650      115              120              125
651
652  TTC CTG GCT TCT GTG AGC ACC GTG CTG ACC TCC AAA TAC CGT TAAGCTGGAG      677
653
654  Phe Leu Ala Ser Val Ser Thr Val Leu Thr Ser Lys Tyr Arg
655      130              135              140
656
657
658  CCTCGGTAGC CGTTCCTCCT GCCCGGTCGA CC      709
659
660
661  (2) INFORMATION FOR SEQ ID NO:8:
662
663  (i) SEQUENCE CHARACTERISTICS:

```

RAW SEQUENCE LISTING
PATENT APPLICATION US/07/923,692CDATE: 10/28/93
TIME: 16:29:09

INPUT SET: S818.raw

664 (A) LENGTH: 141 amino acids
665 (B) TYPE: amino acid
666 (D) TOPOLOGY: linear
667
668 (ii) MOLECULE TYPE: protein
669
670
671 (ix) SEQUENCE DESCRIPTION: SEQ ID NO:8:
672
673 Val Leu Ser Pro Ala Asp Lys Thr Asn Val Lys Ala Ala Trp Gly Lys
674 1 5 10 15
675
676 Val Gly Ala His Ala Gly Glu Tyr Gly Ala Glu Ala Leu Glu Arg Met
677 20 25 30
678
679 Phe Leu Ser Phe Pro Thr Thr Lys Thr Tyr Phe Pro His Phe Asp Leu
680 35 40 45
681
682 Ser His Gly Ser Ala Gln Val Lys Gly His Gly Lys Lys Val Ala Asp
683 50 55 60
684
685 Ala Leu Thr Asn Ala Val Ala His Val Asp Asp Met Pro Asn Ala Leu
686 65 70 75 80
687
688 Ser Ala Leu Ser Asp Leu His Ala His Lys Leu Arg Val Asp Pro Val
689 85 90 95
690
691 Asn Phe Lys Leu Leu Ser His Cys Leu Leu Val Thr Leu Ala Ala His
692 100 105 110
693
694 Leu Pro Ala Glu Phe Thr Pro Ala Val His Ala Ser Leu Asp Lys Phe
695 115 120 125
696
697 Leu Ala Ser Val Ser Thr Val Leu Thr Ser Lys Tyr Arg
698 130 135 140
699
700

(2) INFORMATION FOR SEQ ID NO:9:

701
702
703 (i) SEQUENCE CHARACTERISTICS:
704 (A) LENGTH: 743 base pairs
705 (B) TYPE: nucleic acid
706 (C) STRANDEDNESS: single
707 (D) TOPOLOGY: linear
708
709 (ii) MOLECULE TYPE: cDNA to mRNA
710
711 (iii) HYPOTHETICAL: NO
712
713 (iv) ANTI-SENSE: NO
714

RAW SEQUENCE LISTING PATENT APPLICATION US/07/923,692C

DATE: 10/28/93
TIME: 16:29:16

INPUT SET: S818.raw

```

715 (vi) ORIGINAL SOURCE:
716 (A) ORGANISM: Homo sapiens
717
718 (vii) IMMEDIATE SOURCE:
719 (B) CLONE: beta-hemoglobin
720
721 (ix) FEATURE:
722 (A) NAME/KEY: transit_peptide (B)
723 LOCATION: 26..241
724 (B) LOCATION: 26..241
725
726 (ix) FEATURE:
727 (A) NAME/KEY: CDS
728 (B) LOCATION: 245..685
729
730 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:
731
732 CTCGAGGGGA TCTGATCTTT CAAGAATGGC ACAAATTAAC AACATGGCAC AAGGGATACA 60
733
734 AACCCCTTAAT CCCAATTCCA ATTTCCATAA ACCCCAAGTT CCTAAATCTT CAAGTTTTCT 120
735
736 TGTTTTTTGGA TCTAAAAAAC TGAAAAATTC AGCAAATTCCT ATGTTGGTTT TGAAAAAAGA 180
737
738 TTCAATTTTT ATGCAAAAGT TTTGTTCCCTT TAGGATTTC A GCAGGTGGTA GAGTTTCTTG 240
739
740 GATG GTG CAC CTG ACT CCT GAG GAG AAG TCT GCC GTT ACT GCC CTG TGG 289
741
742 Val His Leu Thr Pro Glu Glu Lys Ser Ala Val Thr Ala Leu Trp
743 1 5 10 15
744
745 GGC AAG GTG AAC GTG GAT GAA GTT GGT GGT GAG GCC CTG GGC AGG CTG 337
746
747 Gly Lys Val Asn Val Asp Glu Val Gly Gly Glu Ala Leu Gly Arg Leu
748 20 25 30
749
750 CTG GTG GTC TAC CCT TGG ACC CAG AGG TTC TTT GAG TCC TTT GGG GAT 385
751
752 Leu Val Val Tyr Pro Trp Thr Gln Arg Phe Phe Glu Ser Phe Gly Asp
753 35 40 45
754
755 CTG TCC ACT CCT GAT GCT GTT ATG GGC AAC CCT AAG GTG AAG GCT CAT 433
756
757 Leu Ser Thr Pro Asp Ala Val Met Gly Asn Pro Lys Val Lys Ala His
758 50 55 60
759
760 GGC AAG AAA GTG CTG GGT GCC TTT AGT GAT GGC CTG GCT CAC CTG GAC 481
761
762 Gly Lys Lys Val Leu Gly Ala Phe Ser Asp Gly Leu Ala His Leu Asp
763 65 70 75
764
765 AAC CTC AAG GGC ACC TTT GCC ACC CTG AGT GAG CTG CAC TGT GAC AAG 529

```


RAW SEQUENCE LISTING PATENT APPLICATION US/07/923,692C

DATE: 10/28/93
TIME: 16:29:23

INPUT SET: S818.raw

```

766
767 Asn Leu Lys Gly Thr Phe Ala Thr Leu Ser Glu Leu His Cys Asp Lys
768 80 85 90 95
769
770 CTG CAC GTG GAT CCT GAG AGC TTC AGG CTC CTA GGC AAC GTG CTG GTC 577
771
772 Leu His Val Asp Pro Glu Ser Phe Arg Leu Leu Gly Asn Val Leu Val
773 100 105 110
774
775 TGT GTG CTG GCG CAT CAC TTT GGC AAA GAA TTC ACC CCA CCA GTG CAG 625
776
777 Cys Val Leu Ala His His Phe Gly Lys Glu Phe Thr Pro Pro Val Gln
778 115 120 125
779
780 GCT GCC TAT CAG AAA GTG GTG GCT GGT GTG GCT AAT GCC CTG GCC CAC 673
781
782 Ala Ala Tyr Gln Lys Val Val Ala Gly Val Ala Asn Ala Leu Ala His
783 130 135 140
784
785 AAG TAT CAC TAAGCTCGCT TTCTTGCTGT CCAATTTCTA TTAAAGGTTC 722
786
787 Lys Tyr His
788 145
789
790 CTTTGTGGGG TCGAGGTCGA C 743
791
792
793
794 (2) INFORMATION FOR SEQ ID NO: 10:
795
796 (i) SEQUENCE CHARACTERISTICS:
797 (A) LENGTH: 146 amino acids
798 (B) TYPE: amino acid
799 (D) TOPOLOGY: linear
800 (ii) MOLECULE TYPE: protein
801
802 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:
803
804 Val His Leu Thr Pro Glu Glu Lys Ser Ala Val Thr Ala Leu Trp Gly
805 1 5 10 15
806
807 Lys Val Asn Val Asp Glu Val Gly Gly Glu Ala Leu Gly Arg Leu Leu
808 20 25 30
809
810 Val Val Tyr Pro Trp Thr Gln Arg Phe Phe Glu Ser Phe Gly Asp Leu
811 35 40 45
812
813 Ser Thr Pro Asp Ala Val Met Gly Asn Pro Lys Val Lys Ala His Gly
814 50 55 60
815
816 Lys Lys Val Leu Gly Ala Phe Ser Asp Gly Leu Ala His Leu Asp Asn

```

RAW SEQUENCE LISTING PATENT APPLICATION US/07/923,692C

 DATE: 10/28/93
 TIME: 16:29:30

INPUT SET: S818.raw

```

817      65              70              75              80
818
819  Leu Lys Gly Thr Phe Ala Thr Leu Ser Glu Leu His Cys Asp Lys Leu
820              85              90              95
821
822  His Val Asp Pro Glu Ser Phe Arg Leu Leu Gly Asn Val Leu Val Cys
823              100              105              110
824
825  Val Leu Ala His His Phe Gly Lys Glu Phe Thr Pro Pro Val Gln Ala
826              115              120              125
827
828  Ala Tyr Gln Lys Val Val Ala Gly Val Ala Asn Ala Leu Ala His Lys
829      130              135              140
830
831  Tyr His
832  145
833
834
835  (2)  INFORMATION FOR SEQ ID NO:11:
836
837      (i)      SEQUENCE CHARACTERISTICS:
838              (A)  LENGTH: 17 amino acids
839              (B)  TYPE: amino acid
840              (D)  TOPOLOGY: linear
841
842      (ii)     MOLECULE TYPE: peptide
843
844      (v)      FRAGMENT TYPE: N-terminal
845
846      (vi)     ORIGINAL SOURCE:
847              (A)  ORGANISM: alkalophilic Bacillus sp.
848              (B)  STRAIN: 38-2
849
850      (vii)    IMMEDIATE SOURCE:
851              (B)  CLONE: beta-cyclodextrin
852
853      (xi)     SEQUENCE DESCRIPTION: SEQ ID NO: 11:
854
855  Ala Pro Asp Thr Ser Val Ser Asn Lys Gln Asn Phe Ser Thr Asp Val
856  1              5              10              15
857
858  Ile
859

```

PAGE: 1

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION US/07/923,692C

DATE: 10/28/93
TIME: 16:29:36

INPUT SET: S818.raw

Line	Error	Original Text
31	Wrong application Serial Number	(A) APPLICATION NUMBER: US 923,692

PAGE: 1

SEQUENCE MISSING ITEM REPORT
PATENT APPLICATION US/07/923,692C

DATE: 10/28/93
TIME: 16:29:36

INPUT SET: S818.raw

COUNTRY
PRIOR APPLICATION DATA More Identifiers Found Than MAX Allowed

PAGE: 1

SEQUENCE CORRECTION REPORT
PATENT APPLICATION US/07/923,692C

DATE: 10/28/93
TIME: 16:29:37

INPUT SET: S818.raw

Line	Original Text	Corrected Text
------	---------------	----------------